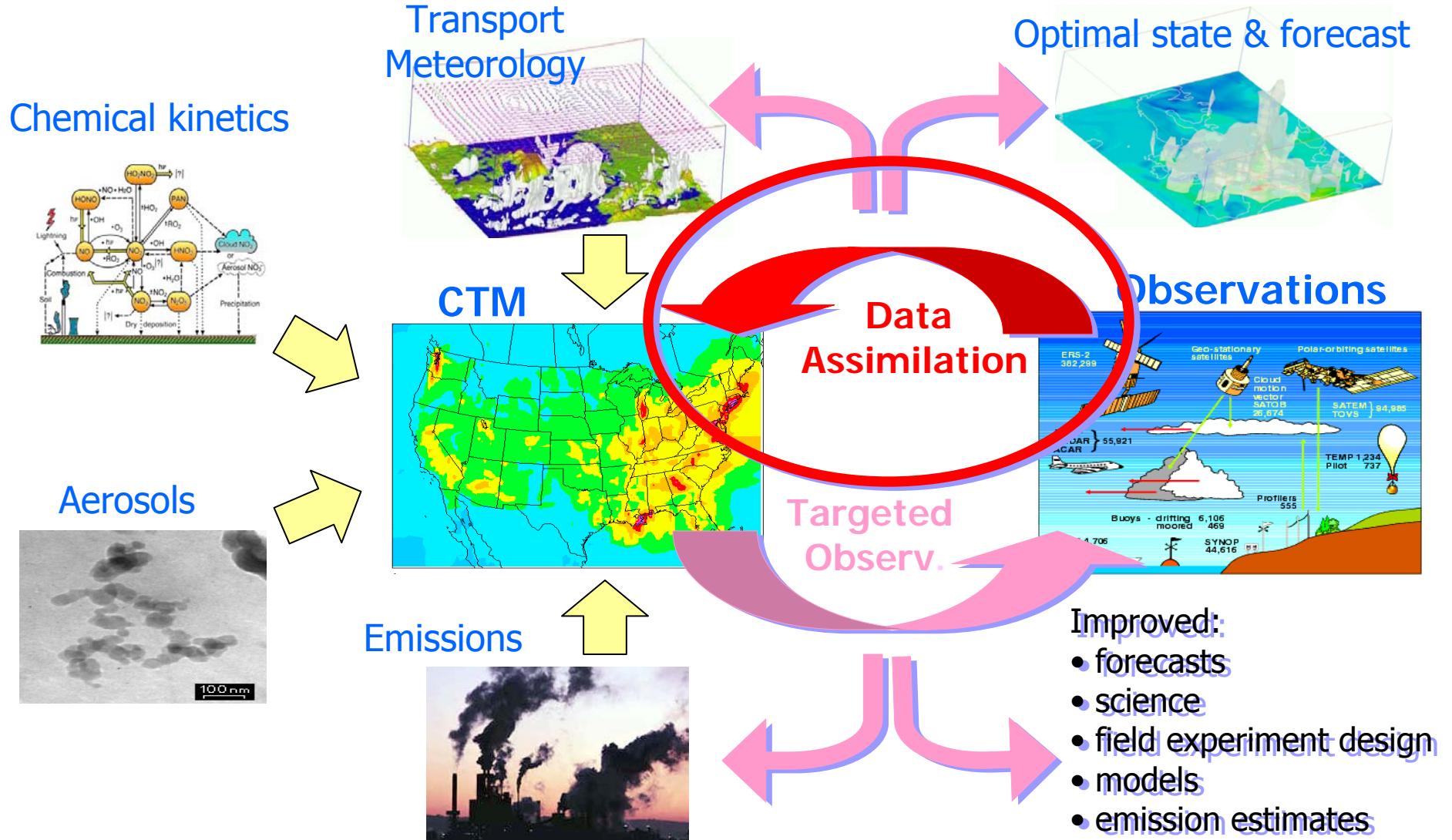


Computational Tools for Data Assimilation in Atmospheric CTMs

A. Sandu, H. Cheng, E.M. Constantinescu, W.
Liao, Z. Liu, P. Miehe, K. Singh, L. Zhang
(Virginia Tech)

U. Iowa, Caltech
NSF, NASA, NOAA

Information feedback loops between CTMs and observations: data assimilation and targeted meas.



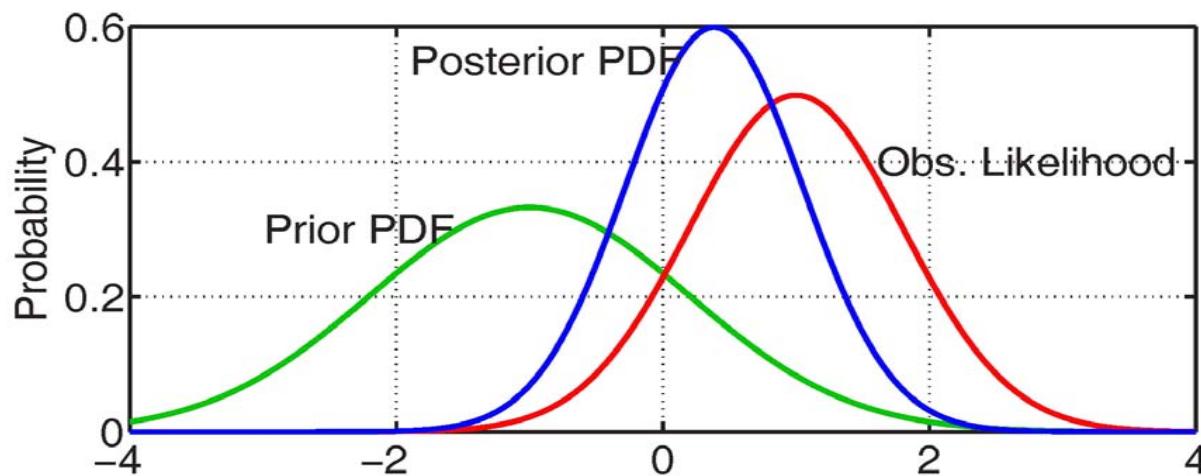
Best estimate of the state is obtained by combining three sources of information

Model (encapsulating knowledge on the physics, chemistry, thermodynamics, etc)

Background (encapsulating best a-priori knowledge of the state) $y^0 \sim N(y^b, B)$

Observations (new information about reality) $z_{\text{obs}} = h(y) + \varepsilon_{\text{obs}}, \varepsilon_{\text{obs}} \sim N(0, R)$

$$\text{Bayes: } P[y^k \mid z_{\text{obs}}^k \dots z_{\text{obs}}^0] = \frac{P[z_{\text{obs}}^k \mid y^k] \cdot P[y^k \mid z_{\text{obs}}^{k-1} \dots z_{\text{obs}}^0]}{\int P[z_{\text{obs}}^k \mid y] \cdot P[y \mid z_{\text{obs}}^{k-1} \dots z_{\text{obs}}^0] dy} = \frac{P_R[\varepsilon_{\text{obs}}^k] \cdot P_B[y^k]}{\dots}$$



Methods:

4D-Var

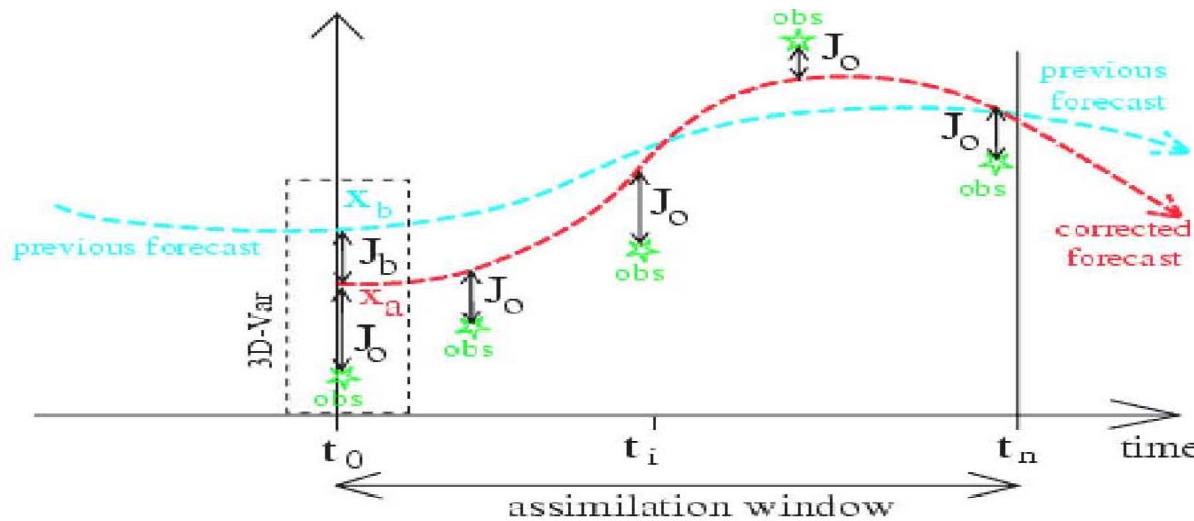
EnKF

[Picture from
J.L. Anderson]

In the 4D-Var approach D.A. is formulated as a PDE-constrained optimization problem (gradient-based)

$$\min_{\mathbf{y}^0} \psi(\mathbf{y}^0) = \frac{1}{2} (\mathbf{y}^0 - \mathbf{y}^b)^T \mathbf{B}^{-1} (\mathbf{y}^0 - \mathbf{y}^b) + \frac{1}{2} \sum_{k=1}^N (\mathbf{H}^k \mathbf{y}^k - \mathbf{z}_{obs}^k)^T \mathbf{R}_k^{-1} (\mathbf{H}^k \mathbf{y}^k - \mathbf{z}_{obs}^k)$$

subject to $\mathbf{y}^k = M(t^{k-1}, \mathbf{y}^{k-1}, \mathbf{p}), \quad k = 1, 2, \dots$



MAP estimate of \mathbf{y}^0
(Nonl. Red. Q-Newt.)

[Picture from
F. Bouttier, P. Courtier]

Gradient: $\lambda^0 = \nabla_{\mathbf{y}^0} \psi = \left(\frac{\partial \psi}{\partial \mathbf{y}^0} \right)^T = \mathbf{B}^{-1} (\mathbf{y}^0 - \mathbf{y}^b) + \sum_{k=1}^N \left(\frac{\partial \mathbf{y}^k}{\partial \mathbf{y}^0} \right)^T (\mathbf{H}^k)^T \mathbf{R}_k^{-1} (\mathbf{H}^k \mathbf{y}^k - \mathbf{z}_{obs}^k)$

Continuous and discrete adjoints of mass balance equations lead to different computational models

$$\nabla_{\mathbf{y}^0} \psi = \cdots + \sum_{k=1}^N \left(\partial \mathbf{y}^k / \partial \mathbf{y}^0 \right)^T \left(\mathbf{H}^k \right)^T \mathbf{R}_k^{-1} \left(\mathbf{H}^k \mathbf{y}^k - \mathbf{z}_{obs}^k \right)$$

Continuous forward model

$$\begin{aligned} \frac{dC_i}{dt} &= -\vec{u} \cdot \nabla C_i + \frac{1}{\rho} \nabla (\rho K \cdot \nabla C_i) + \frac{1}{\rho} f_i(\rho C) + E_i \\ C_i(t^0, x) &= C_i^0(x), \quad t^0 \leq t \leq t^F \\ C_i(t, x) &= C_i^{IN}(t, x) \quad \text{on } \Gamma^{IN} \\ K \frac{\partial C_i}{\partial n} &= 0 \quad \text{on } \Gamma^{OUT} \\ K \frac{\partial C_i}{\partial n} &= V_i^{DEP} C_i - Q_i \quad \text{on } \Gamma^{GROUND} \end{aligned}$$

adj

Continuous adjoint model

$$\begin{aligned} \frac{d\lambda_i}{dt} &= -\nabla \cdot (\vec{u} \lambda_i) - \nabla \cdot \left(\rho K \cdot \nabla \frac{\lambda_i}{\rho} \right) - (F^T(\rho C) \cdot \lambda)_i - \phi_i \\ \lambda_i(t^F, x) &= \lambda_i^F(x), \quad t^F \geq t \geq t^0 \\ \lambda_i(t, x) &= 0 \quad \text{on } \Gamma^{IN} \\ \vec{u} \lambda_i + \rho K \frac{\partial(\lambda_i/\rho)}{\partial \vec{n}} &= 0 \quad \text{on } \Gamma^{OUT} \\ \rho K \frac{\partial(\lambda_i/\rho)}{\partial \vec{n}} &= V_i^{DEP} \lambda_i \quad \text{on } \Gamma^{GROUND} \end{aligned}$$

discr

Discrete forward model

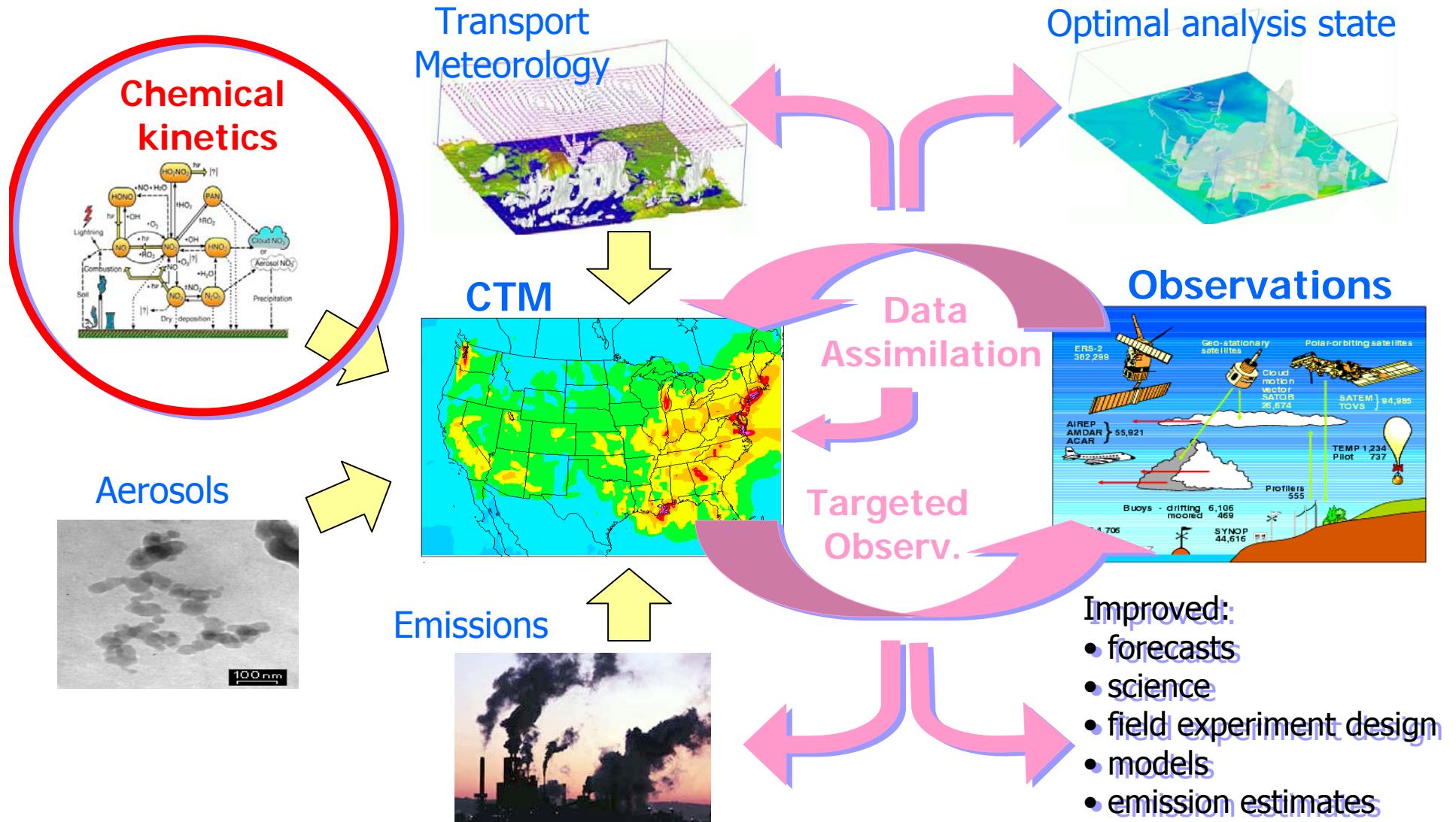
$$\begin{aligned} C^{k+1} &= N_{[t, t+\Delta t]} \circ C^k \\ N_{[t, t+\Delta t]} &= T_{HOR}^{\Delta t} \circ T_{VERT}^{\Delta t} \circ R_{CHEM}^{\Delta t} \circ T_{VERT}^{\Delta t} \circ T_{HOR}^{\Delta t} \end{aligned}$$

adj

Computational adjoint model

$$\begin{aligned} \lambda^k &= N^*_{[t, t+\Delta t]} \circ \lambda^{k+1} + \phi^{k+1} \\ N^*_{[t, t+\Delta t]} &= (T_{HOR}^{\Delta t})^* \circ (T_{VERT}^{\Delta t})^* \circ (R_{CHEM}^{\Delta t})^* \circ (T_{VERT}^{\Delta t})^* \circ (T_{HOR}^{\Delta t})^* \end{aligned}$$

Adjoints of stiff chemical kinetics: formulation, challenges, and automatic implementation



KPP automatically generates simulation and direct/adjoint sensitivity code for chemistry

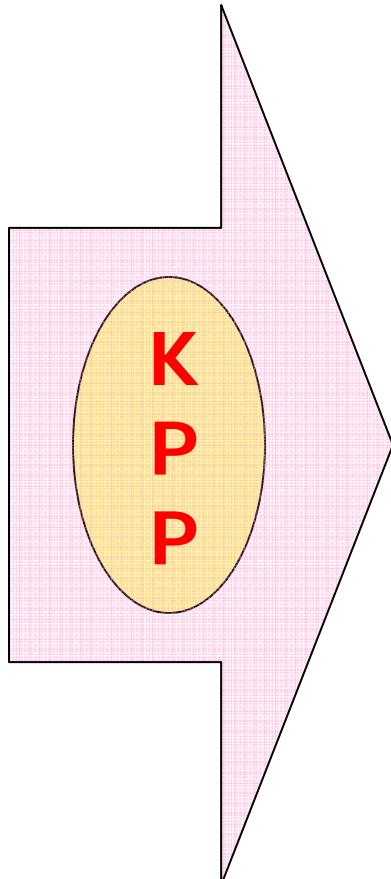
Chemical mechanism

```
#INCLUDE atoms

#DEFVAR
O = O; O1D = O;
O3 = O + O + O;
NO = N + O;
NO2 = N + O + O;

#DEFFIX
O2 = O + O; M = ignore;

#EQUATIONS { Small Stratospheric }
O2 + hv = 2O : 2.6E-10*S;
O + O2 = O3 : 8.0E-17;
O3 + hv = O + O2 : 6.1E-04*S;
O + O3 = 2O2 : 1.5E-15;
O3 + hv = O1D + O2 : 1.0E-03*S;
O1D + M = O + M : 7.1E-11;
O1D + O3 = 2O2 : 1.2E-10;
NO + O3 = NO2 + O2 : 6.0E-15;
NO2 + O = NO + O2 : 1.0E-11;
NO2 + hv = NO + O : 1.2E-02*S;
```



Simulation code

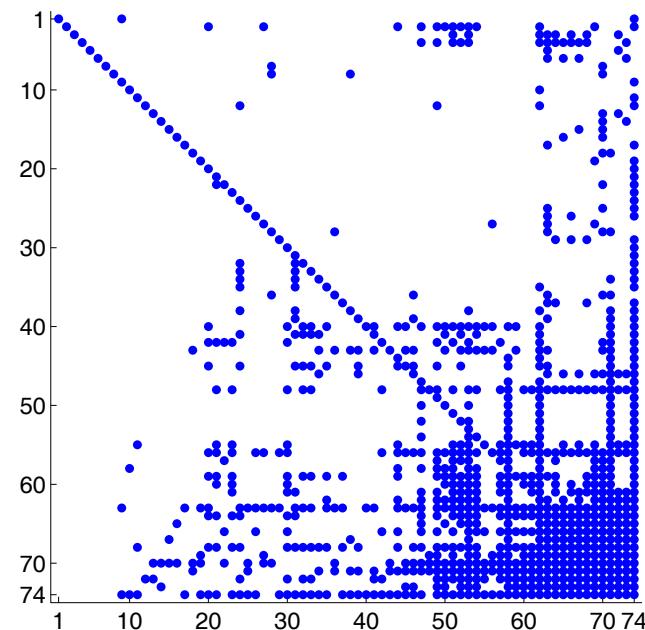
```
SUBROUTINE FunVar (V, F, RCT, DV)
INCLUDE 'small.h'
REAL*8 V(NVAR), F(NFIX)
REAL*8 RCT(NREACT), DV(NVAR)
C A - rate for each equation
REAL*8 A(NREACT)
C Computation of equation rates
A(1) = RCT(1)*F(2)
A(2) = RCT(2)*V(2)*F(2)
A(3) = RCT(3)*V(3)
A(4) = RCT(4)*V(2)*V(3)
A(5) = RCT(5)*V(3)
A(6) = RCT(6)*V(1)*F(1)
A(7) = RCT(7)*V(1)*V(3)
A(8) = RCT(8)*V(3)*V(4)
A(9) = RCT(9)*V(2)*V(5)
A(10) = RCT(10)*V(5)
C Aggregate function
DV(1) = A(5)-A(6)-A(7)
DV(2) = 2*A(1)-A(2)+A(3)-A(4)+A(6)-&A(9)+A(10)
DV(3) = A(2)-A(3)-A(4)-A(5)-A(7)-A(8)
DV(4) = -A(8)+A(9)+A(10)
DV(5) = A(8)-A(9)-A(10)
END
```

[Damian et.al., 1996; Sandu et.al., 2002]

Sparse Jacobians , Hessians, and sparse linear algebra routines are automatically generated by KPP

#JACOBIAN [ON | OFF | SPARSE]

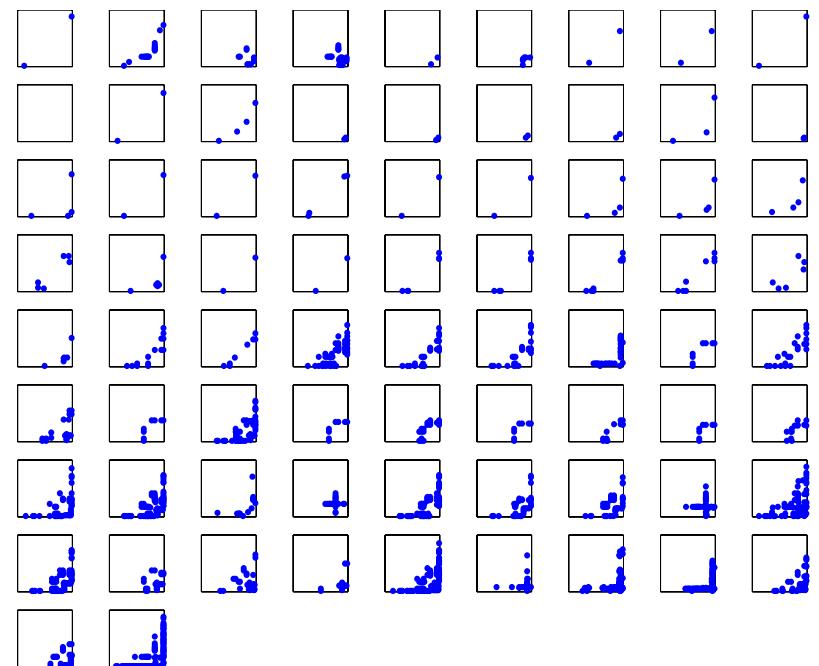
JacVar(...), JacVar[TR]_SP_Vec(...)
KppDecomp(...), KppSolve[TR](...)



SAPRC-99
79 spc./211 react. NZ=839, NZLU=920

#HESSIAN [ON | OFF]

HessVar(...), HessVar[TR]_Vec(...)



SAPRC-99.
NZ = 848x2 (0.2%)

Runge-Kutta methods and their adjoints are well suited for inverse chemical kinetic problems

RK Method

$$\mathbf{y}^{n+1} = \mathbf{y}^n + h \sum_{i=1}^s b_i \mathbf{f}(\mathbf{Y}^i), \quad \mathbf{Y}^i = \mathbf{y}^n + h \sum_{j=1}^s a_{i,j} \mathbf{f}(\mathbf{Y}^j)$$

Continuous
Adjoint

$$\lambda^n = \lambda^{n+1} + h \sum_{i=1}^s b_i \mathbf{J}^T(\mathbf{y}^{n+1} - \mathbf{c}_i h) \cdot \Lambda^i, \quad \Lambda^i = \lambda^{n+1} + h \sum_{j=1}^s a_{i,j} \mathbf{J}^T(\mathbf{y}^{n+1} - \mathbf{c}_i h) \cdot \Lambda^j$$

Discrete
Adjoint
[Hager, 2000]

$$\lambda^n = \lambda^{n+1} + \sum_{i=1}^s \theta^i, \quad \theta^i = h \mathbf{J}^T(\mathbf{Y}^i) \cdot \left[b_i \lambda^{n+1} + \sum_{j=1}^s a_{j,i} \theta^j \right]$$

Consistency: The discrete adjoint of RK method of order p is an order p discretization of the adjoint ODE. (Proof using elementary differentials of transfer functions).

Stiff behavior. For SPP apply RK with invertible coefficient matrix A and $R(\infty) = 0$. If the cost function depends only on the non-stiff variable y then $\lambda_z = 0$ and λ_+ are solved with the same accuracy as the original method, within $O(\varepsilon)$.

[Sandu, 2005]

Second order adjoints provide Hessian-vector products useful in optimization and analysis

$$\boldsymbol{\lambda}^0 = \nabla_{\mathbf{y}^0} \psi \quad \boldsymbol{\sigma}^0 = \left(\nabla_{\mathbf{y}^0, \mathbf{y}^0}^2 \psi \right) \cdot \delta \mathbf{y}^0$$

RK & TLM
Methods
(KPP)

$$\mathbf{y}^{n+1} = \mathbf{y}^n + h \sum_{i=1}^s b_i \mathbf{f}(\mathbf{Y}^i), \quad \mathbf{Y}^i = \mathbf{y}^n + h \sum_{j=1}^s a_{i,j} \mathbf{f}(\mathbf{Y}^j)$$

$$\delta \mathbf{y}^{n+1} = \delta \mathbf{y}^n + h \sum_{i=1}^s b_i \mathbf{J}(\mathbf{Y}^i) \delta \mathbf{Y}^i, \quad \delta \mathbf{Y}^i = \delta \mathbf{y}^n + h \sum_{j=1}^s a_{i,j} \mathbf{f}(\mathbf{Y}^j) \delta \mathbf{Y}^j$$

$$\boldsymbol{\lambda}^n = \boldsymbol{\lambda}^{n+1} + \sum_{i=1}^s \mathbf{u}^i + \frac{\partial \Phi}{\partial \mathbf{y}^n}, \quad \mathbf{u}^i = h \mathbf{J}^T(\mathbf{Y}^i) \cdot \left(b_i \boldsymbol{\lambda}^{n+1} + \sum_{j=1}^s a_{j,i} \mathbf{u}^j \right)$$

$$\boldsymbol{\sigma}^n = \boldsymbol{\sigma}^{n+1} + \sum_{i=1}^s \mathbf{w}^i + \frac{\partial^2 \Phi}{\partial^2 \mathbf{y}^n} \delta \mathbf{y}^n,$$

$$\mathbf{w}^i = h \mathbf{J}^T(\mathbf{Y}^i) \cdot \left(b_i \boldsymbol{\sigma}^{n+1} + \sum_{j=1}^s a_{j,i} \mathbf{w}^j \right) + h (\mathbf{H}(\mathbf{Y}^i) \times \delta \mathbf{Y}^i)^T \cdot \left(b_i \boldsymbol{\lambda}^{n+1} + \sum_{j=1}^s a_{j,i} \mathbf{u}^j \right)$$

First and Second
Order RK
Discrete Adjoints
(KPP)

[Sandu et. al., 2007]

Rosenbrock methods and their adjoints are efficiently implemented by KPP

Rosenbrock
Method
 (T_{fwd})

$$\begin{cases} \mathbf{y}^{\mathbf{n+1}} = \mathbf{y}^{\mathbf{n}} + \sum_{j=1}^s m_j \mathbf{k}_j, \quad \mathbf{Y}^i = \mathbf{y}^{\mathbf{n}} + \sum_{j=1}^{i-1} a_{i,j} \mathbf{k}_j \\ \left[\frac{1}{h\gamma} \mathbf{I} - \mathbf{J}^{\mathbf{n}} \right] \cdot \mathbf{k}_i = \mathbf{f}(\mathbf{Y}^i) + \sum_{j=1}^{i-1} \frac{1}{h} c_{i,j} \mathbf{k}_j, \quad 1 \leq i \leq s \end{cases}$$

Continuous
Adjoint
 $(T \approx 1.2 T_{fwd})$

$$\begin{cases} \boldsymbol{\lambda}^{\mathbf{n}} = \boldsymbol{\lambda}^{\mathbf{n+1}} + \sum_{j=1}^s m_j \mathbf{z}_j, \quad \boldsymbol{\Lambda}^i = \boldsymbol{\lambda}^{\mathbf{n+1}} + \sum_{j=1}^{i-1} a_{i,j} \mathbf{z}_j, \quad \mathbf{Y}^i = \mathbf{y}(t^{n+1} - h\alpha_i) \\ \left[\frac{1}{h\gamma} \mathbf{I} - (\mathbf{J}^{\mathbf{n+1}})^T \right] \cdot \mathbf{z}_i = \mathbf{J}^T(\mathbf{Y}^i) \cdot \boldsymbol{\Lambda}^i + \sum_{j=1}^{i-1} \frac{1}{h} c_{i,j} \mathbf{z}_j, \quad 1 \leq i \leq s \end{cases}$$

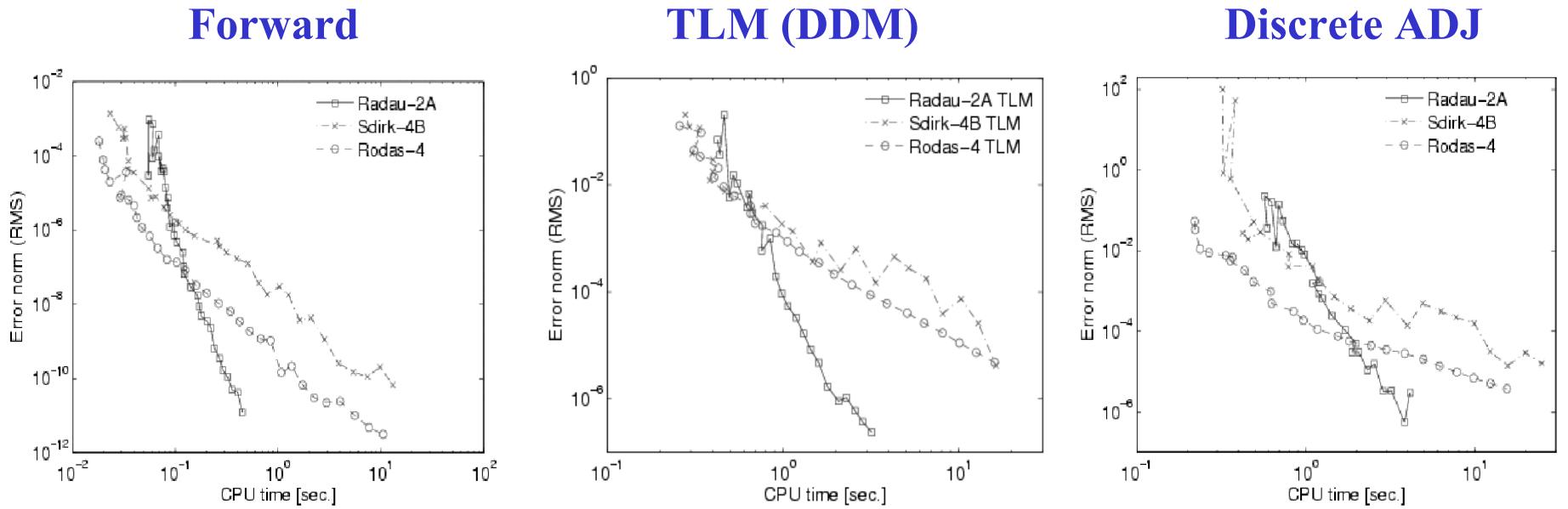
Discrete
Adjoint
 $(T \approx 2.3 T_{fwd})$

$$\begin{cases} \left[\frac{1}{h\gamma} \mathbf{I} - (\mathbf{J}^{\mathbf{n}})^T \right] \cdot \mathbf{u}_i = m_i \boldsymbol{\lambda}^{\mathbf{n+1}} + \sum_{j=i+1}^s (a_{j,i} \mathbf{v}_j + \frac{1}{h} c_{j,i} \mathbf{u}_j), \quad \mathbf{v}_i = \mathbf{J}^T(\mathbf{Y}^i) \cdot \mathbf{u}_i \\ \boldsymbol{\lambda}^{\mathbf{n}} = \boldsymbol{\lambda}^{\mathbf{n+1}} + \sum_{i=1}^s (\mathbf{H}^{\mathbf{n}} \times \mathbf{k}_i)^T \cdot \mathbf{u}_i + \sum_{i=1}^s \mathbf{v}_i \end{cases}$$

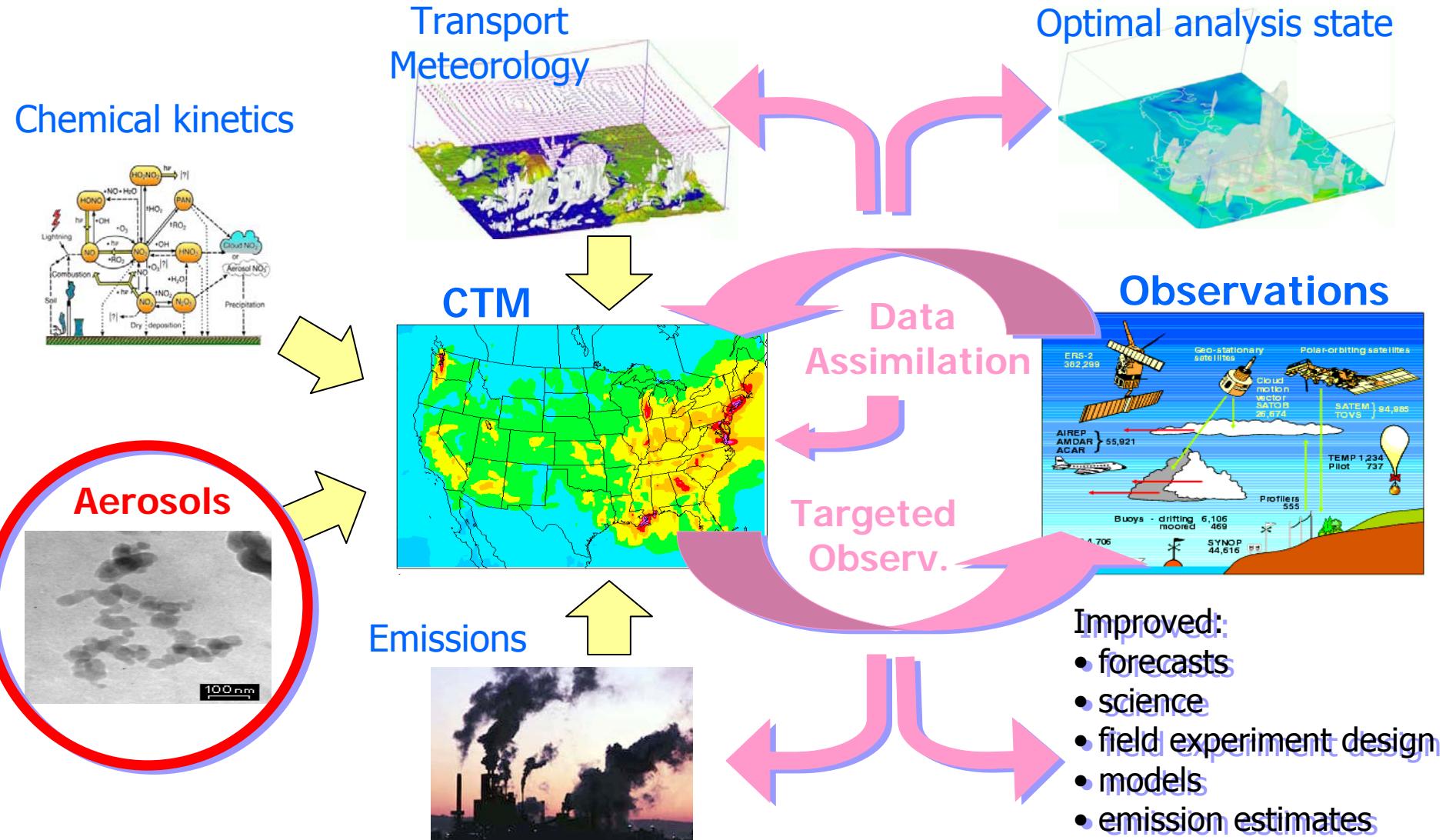
[Sandu et.al., 2002]

Methods available in the KPP and in ADERKDENS libraries

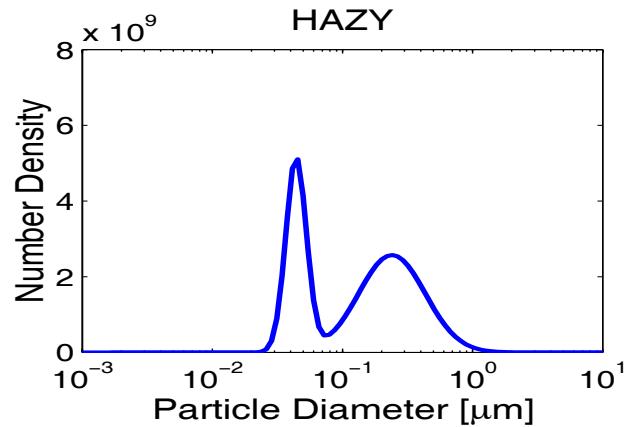
- **FIRK** 3-stage: Radau-2A (ord.5), Radau-1A (ord.5),
Lobatto-3C (ord.4), Gauss (ord.6)
- **SDIRK**: 2a, 2b (2s, ord.2), 3a (3s, ord.2), 4a, 4b (5s, ord.4)
- **Rosenbrock**: Ros2, Ros3, Ros4, Rodas3, Rodas4.
- **ERK**: orders 2-8



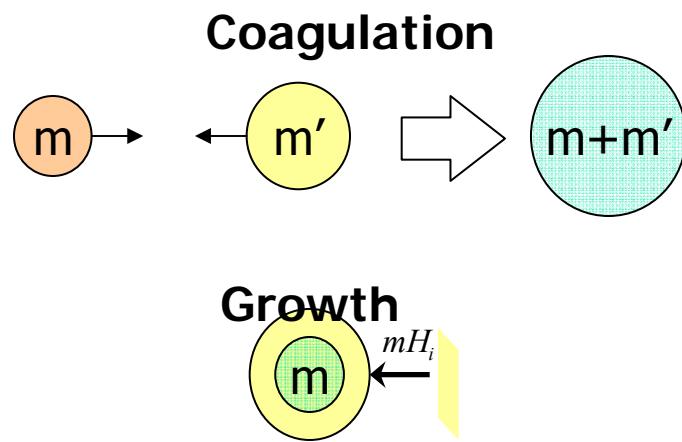
Adjoints for Integral-PDE aerosol dynamic equations: formulation and challenges



Populations of aerosols (particles in the atmosphere) are described by their mass density



Aerosol dynamic equation - IPDE



$$\begin{aligned}
 \frac{\partial q_i}{\partial t} = & \int_0^m \beta(m', m - m') q_i(m', t) \frac{q(m - m', t)}{m - m'} dm' \\
 & - q_i \int_0^\infty \beta(m, m') \frac{q(m', t)}{m'} dm' \\
 & + H_i q - \frac{\partial}{\partial m} (m H_i q_i) + m_i S - L q_i + R_i(q) \\
 q_i(m, t = t^0) = & q_i^0(m), \quad 1 \leq i \leq n, \\
 q_i(m = 0, t) = & 0, \quad q_i(m = \infty, t) = 0.
 \end{aligned}$$

Adjoint aerosol dynamic models are needed to solve inverse problems

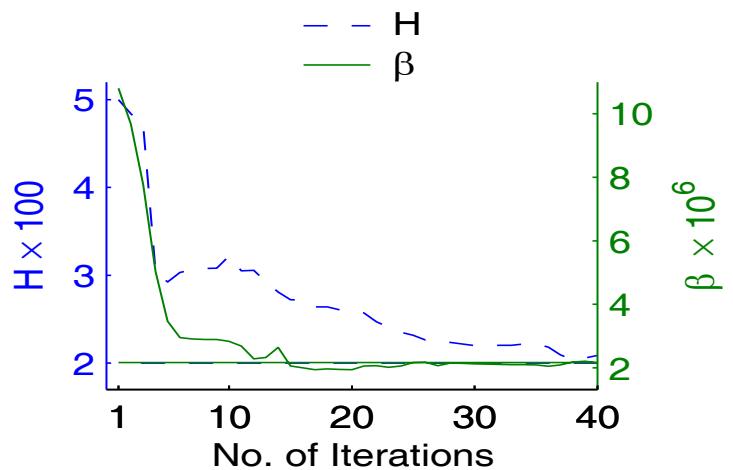
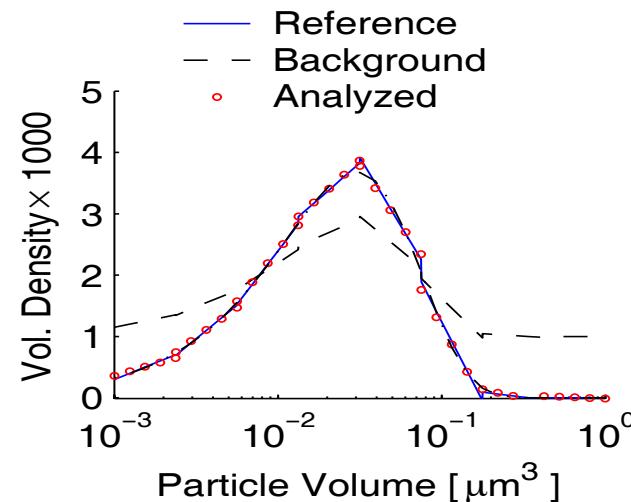
$$\frac{\partial \lambda_i}{\partial t} = - \int_0^{\infty} \beta(m, m') (m')^{-1} [\lambda_i(m+m', t) - \lambda_i(m, t)] q(m', t) dm' + L \lambda_i \quad t_{+}^{k-1} \leq t \leq t_{-}^k$$

Continuous adjoint equation

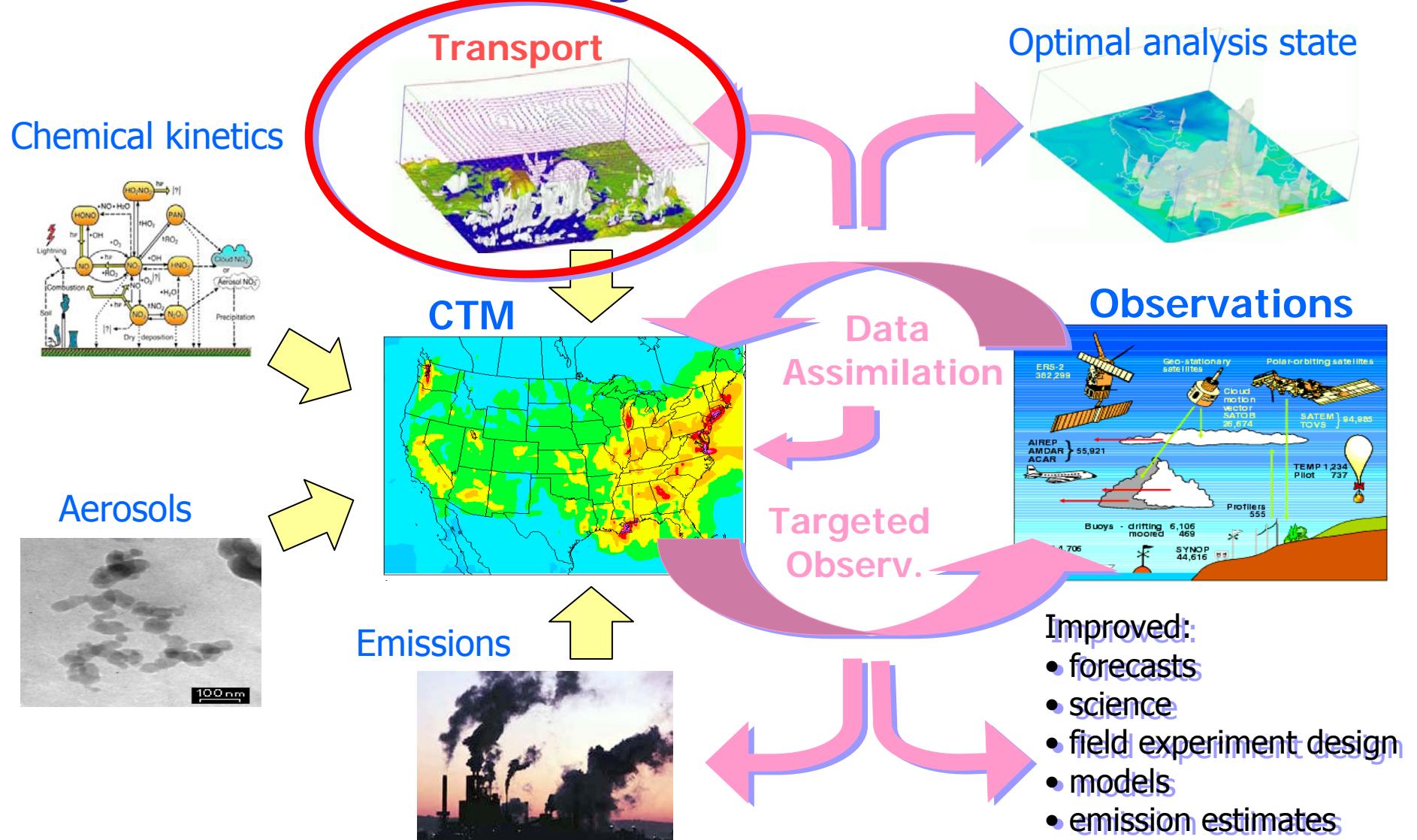
$$\begin{aligned} & - \int_0^{\infty} \beta(m', m) m^{-1} \sum_{j=1}^n [\lambda_j(m+m', t) - \lambda_j(m, t)] q_j(m', t) dm' - \sum_{j=1}^n H_j \lambda_j - m H \frac{\partial \lambda_i}{\partial m} \\ & \lambda_i(m, t^N) = 0, \quad \lambda_i(m, t_{-}^k) = \lambda_i(m, t_{+}^k) + h_{q_i}^T R_k^{-1} (y^k - h(q^k)) \\ & \lambda_i(m, t^0) = \lambda_i(m, t_{+}^0) + p_{q_i}^T B^{-1} (p - p^B), \quad \lambda_i(0, t) = 0. \end{aligned}$$

Observations of density in each bin allow the recovery of initial distribution and of parameters

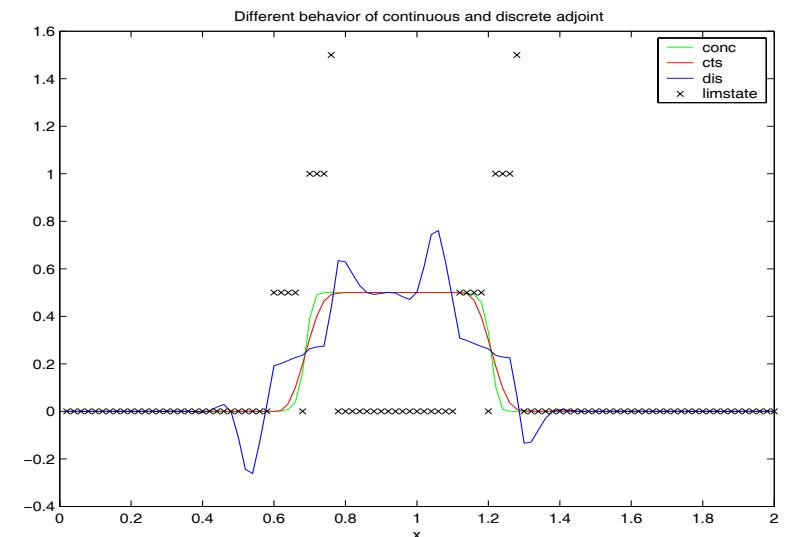
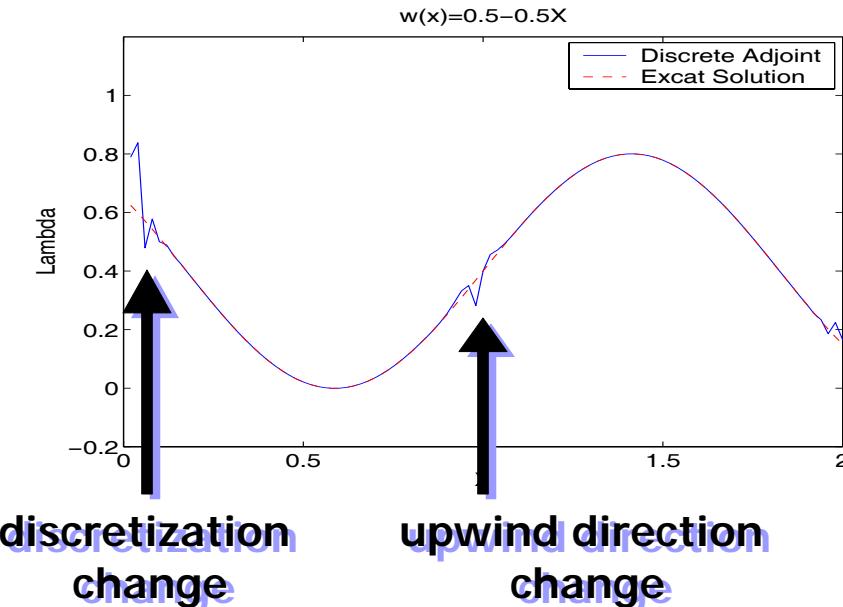
[Henze et. al., 2004;
Sandu et. al., 2005]



Discrete adjoint models for numerical advection: formulation and challenges



Discrete adjoints of advection numerical schemes can become pointwise inconsistent with the adjoint PDE



Change of forward scheme pattern:

- Change of upwinding
- Sources/sinks
- Inflow boundaries scheme

Example: 3rd order upwind FD

Active forward limiters
act as pseudo-sources in adjoint
Example: minmod

[Liu and Sandu, 2005]

AR model of background errors accounts for flow-dependent correlations and is inexpensive

$$\psi(\mathbf{y}^0) = \frac{1}{2} (\mathbf{y}^0 - \mathbf{y}^b)^T \mathbf{B}^{-1} (\mathbf{y}^0 - \mathbf{y}^b) + \dots$$

- Background error repres. considerably impacts the assimilation results
- Typically estimated empirically from multiple model runs (NMC)
- "Correct" mathematical models of background errors are of great interest

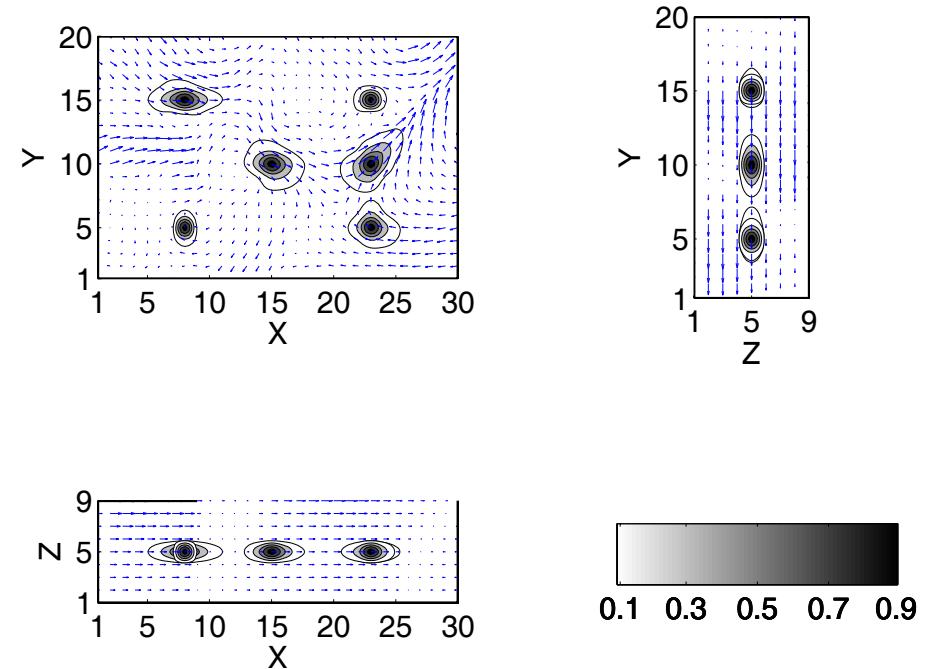
$$\delta\mathbf{y}' = \mathbf{M}'\delta\mathbf{y}$$

$$(\mathbf{I} - \Delta t \mathbf{M}')^N \delta\mathbf{y} = \xi$$

$$\mathbf{B}^{-1} = (\mathbf{I} - \Delta t \mathbf{M}'^*)^N (\mathbf{I} - \Delta t \mathbf{M}')^N$$

- "Monotonic TLM discretization"
- AR model of background errors
- $N\Delta t \approx$ lifetime of the species
- \mathbf{B} is flow dep., cheap, full rank

[Constantinescu et.al., 2005]



The 4D-Var tools have been implemented in parallel adjoint STEM and are being applied to real data

Chemistry: KPP

- Forward: sparse Rosenbrock, RK, LMM
- DDM sensitivity (Rosenbrock, LMM)
- Discrete adjoints: Rosenbrock
- Continuous adjoints: Rosenbrock, RK, LMM

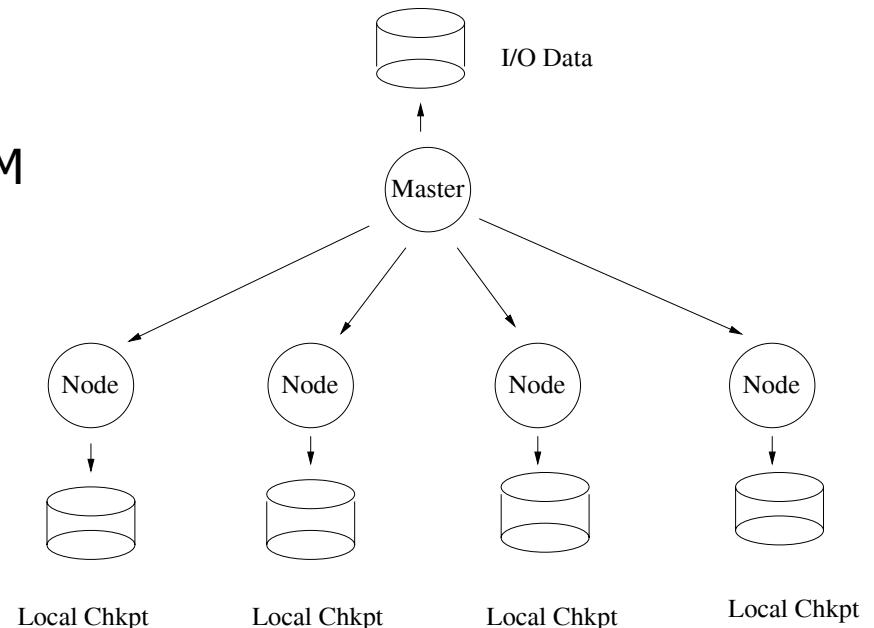
Aerosols: 0-D, not yet 3-D

Transport:

- Forward: upwind FV, FD, FE
- Adjoint for linear upwind FD

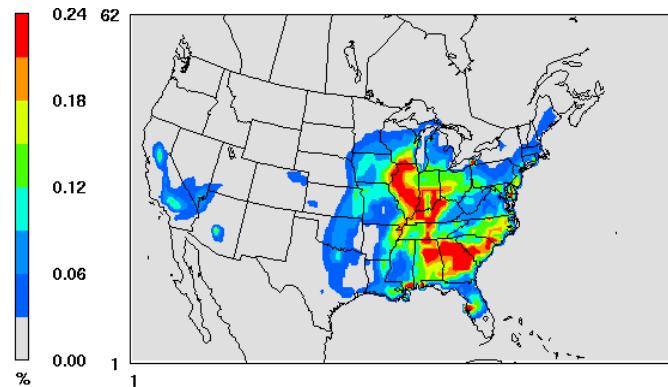
Parallelization: with PAQMSG

Distributed, 2-level checkpointing scheme

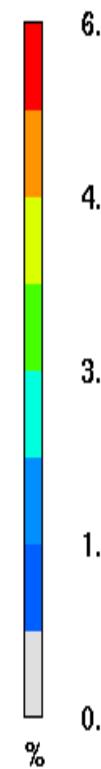
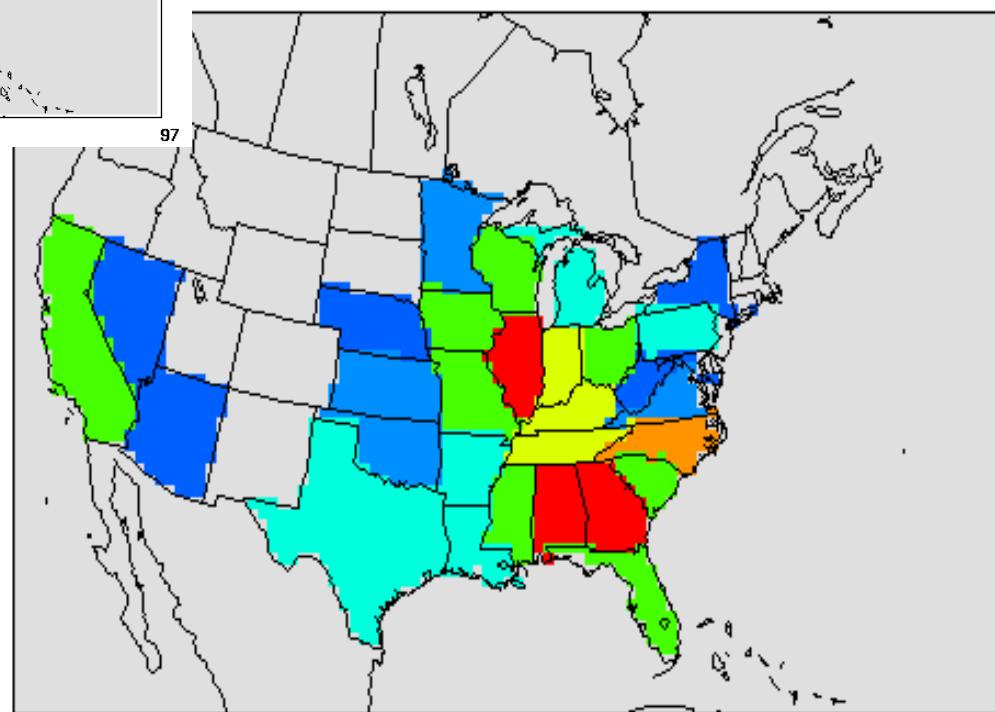


[Sandu et.al., 2003, 2004; Carmichael et. al., 2003, 2004]

Adjoint sensitivity analysis of non-attainment metrics can help guide policy decisions

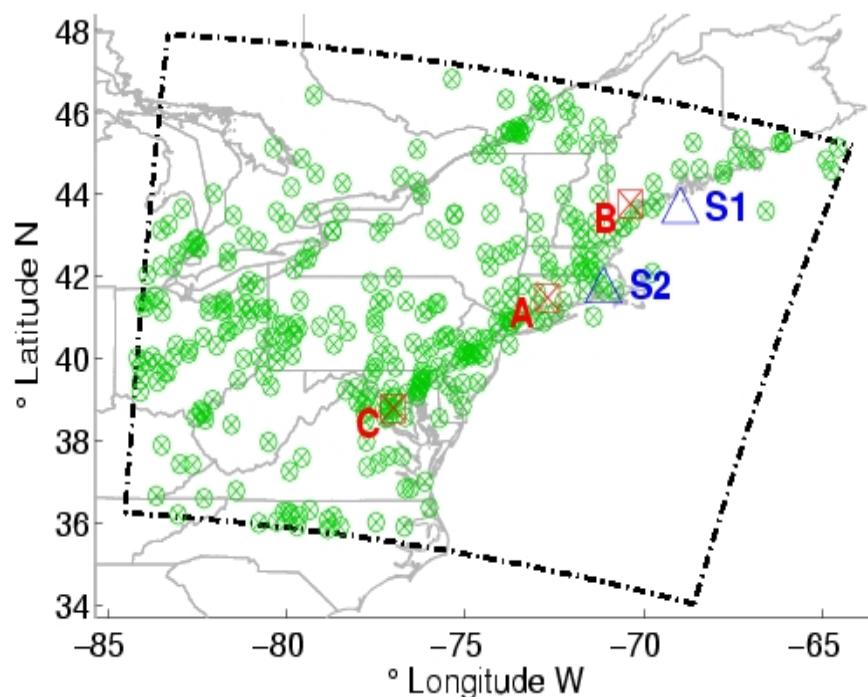


**Estimated contributions by state
to violating U.S. ozone NAAQS
in July 2004**



[Hakami et al., 2005]

Assimilation of ozone data from the ICARTT field campaign in Eastern U.S., July 2004



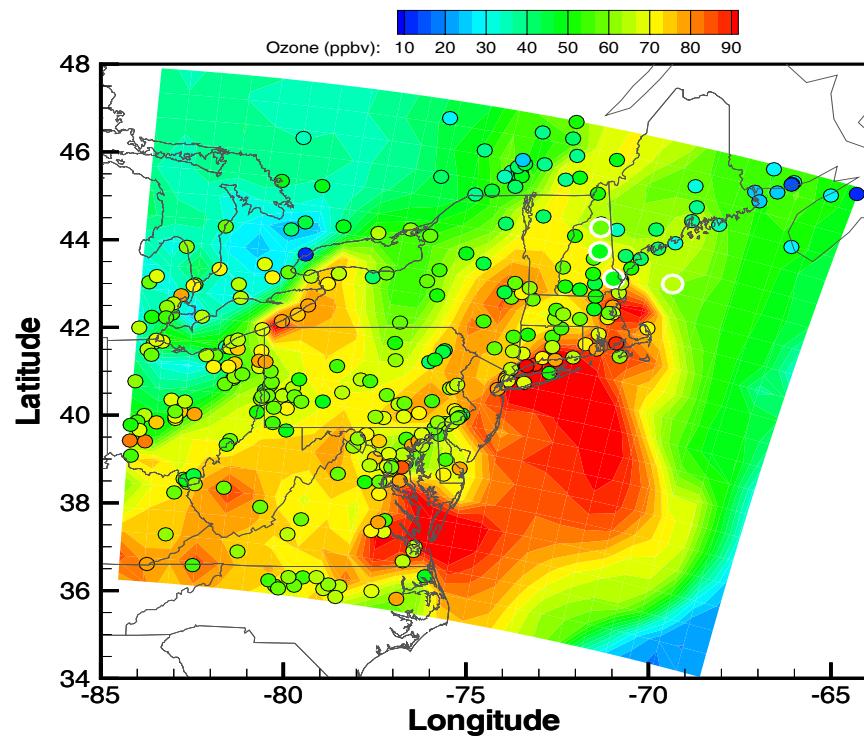
Observations	Description
AIRNOW	EPA surface stations, hourly averaged data used
DC3	Vertical profile of ozone mixing ratio from lidar
MOZ-FN	MOZAIC, Frankfurt-New York flight
MOZ-NF	MOZAIC, New York-Frankfurt flight
P3	NOAA P3-B measurement
AIRMAP	UV SPECTROSCOPY measurement at 4 sites
DC8-In	NASA In Situ Ozone via Nitric Oxide Chemiluminescence
DC8-Li	DC-8 Composite Tropospheric Ozone Cross-Sections
RHODE	Ozonesonde/Radiosonde data from Narragansett, RI
RONBR	Ozonesonde/Radiosonde data from the R/V Ronald H. Brown

[Chai et al., 2006]

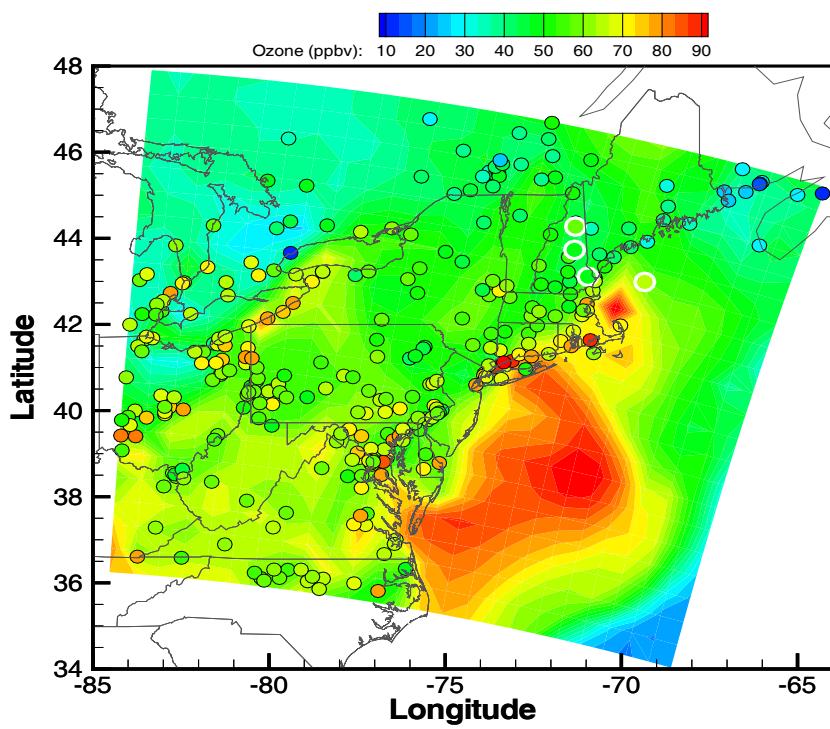
Assimilation adjusts O₃ predictions considerably at 4pm EDT on July 20, 2004

Observations: circles, color coded by O₃ mixing ratio

Surface O₃ (forecast)

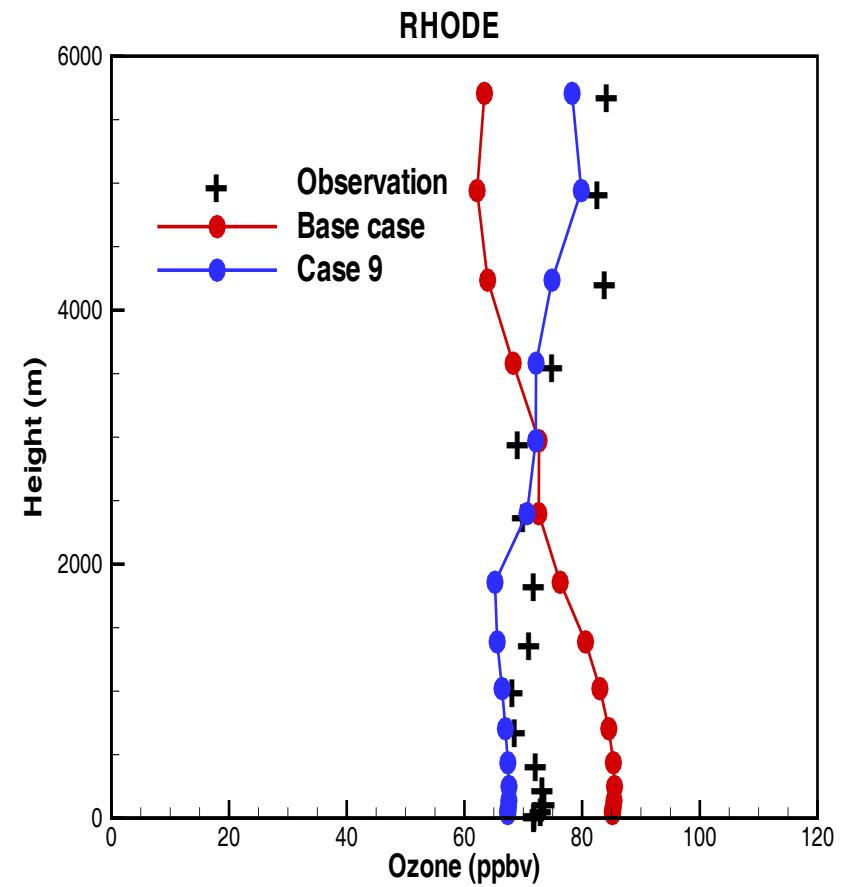
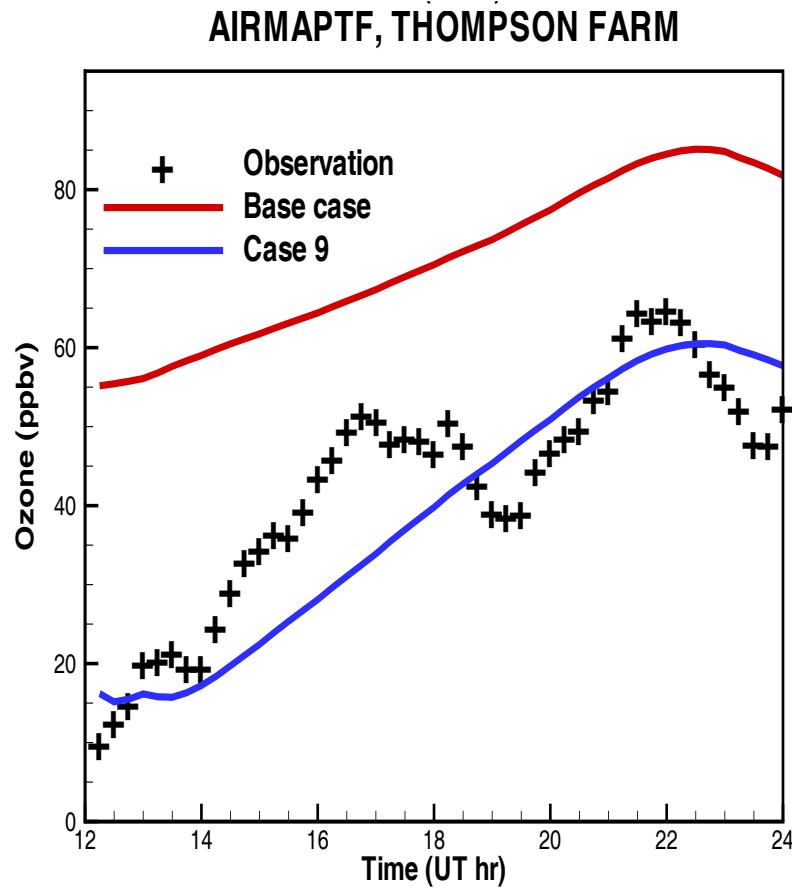


Surface O₃ (analysis)



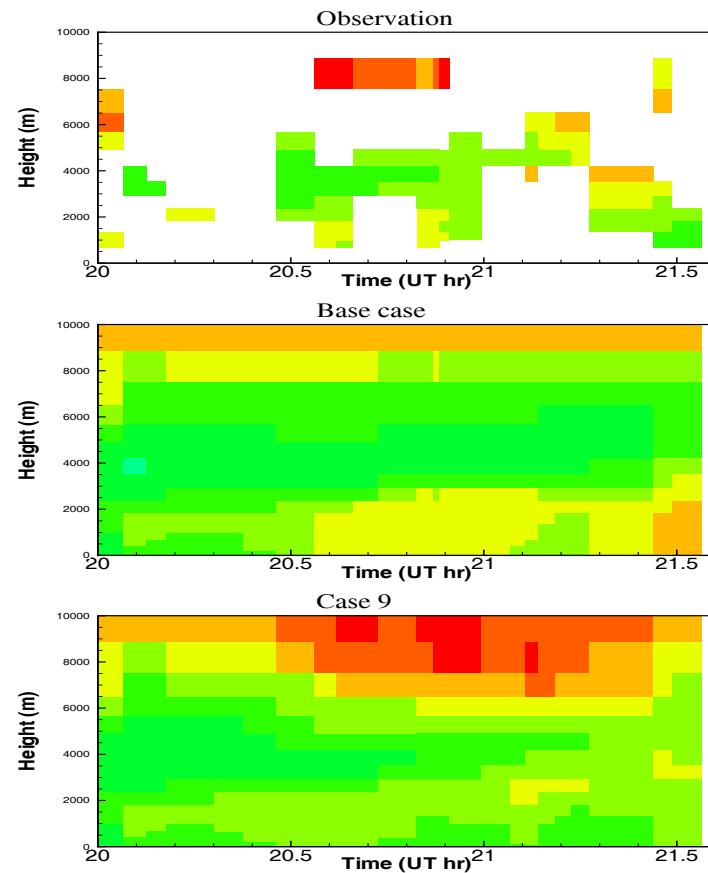
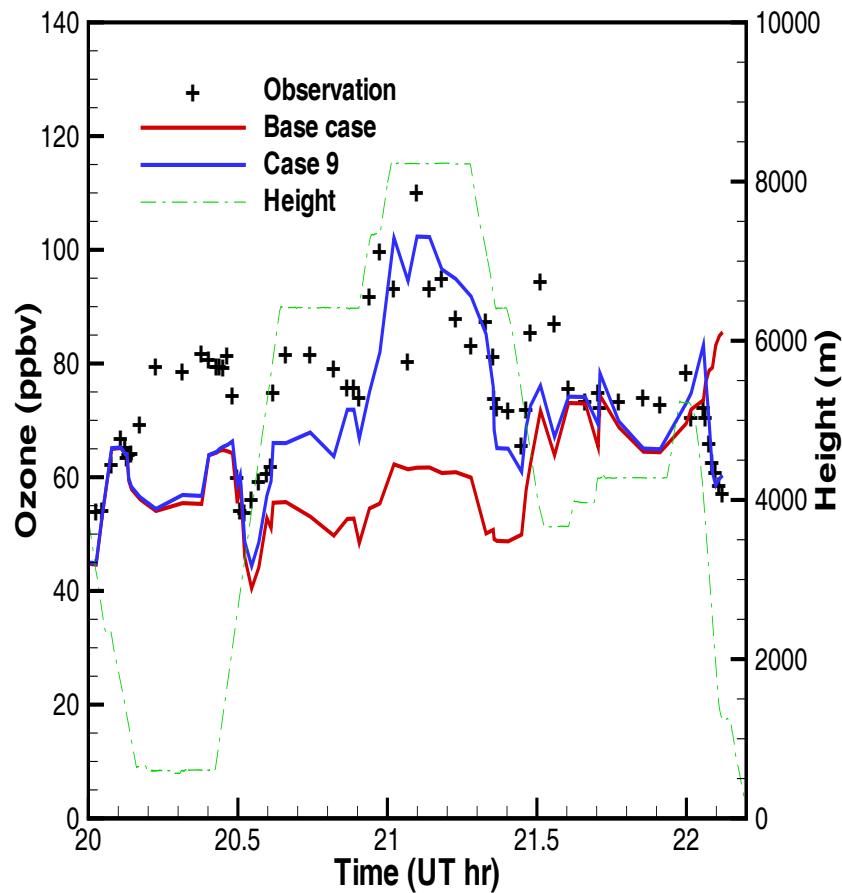
[Chai et al., 2006]

Ozone time series at both ground stations and sondes agree better with observations after assimilation



[Chai et al., 2006]

Assimilation of DC-8 in-situ and lidar observations for July 20, 2004

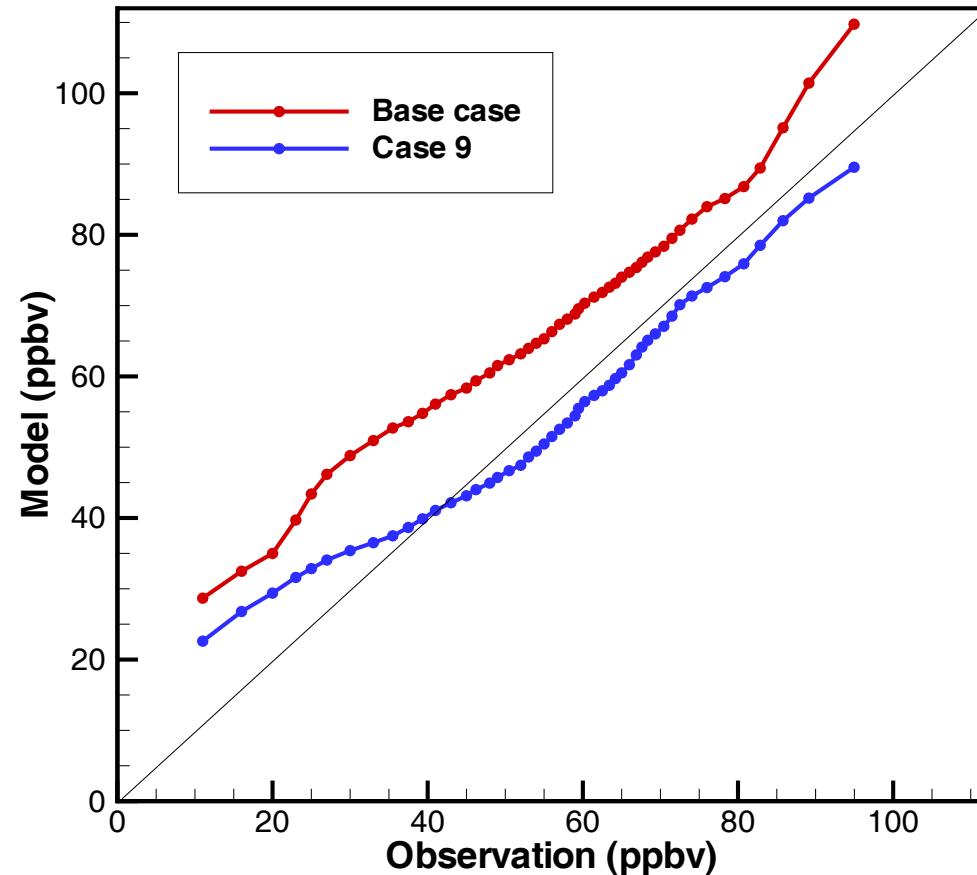


[Chai et al., 2006]

The model results show a considerably improved agreement with observations after data assimilation

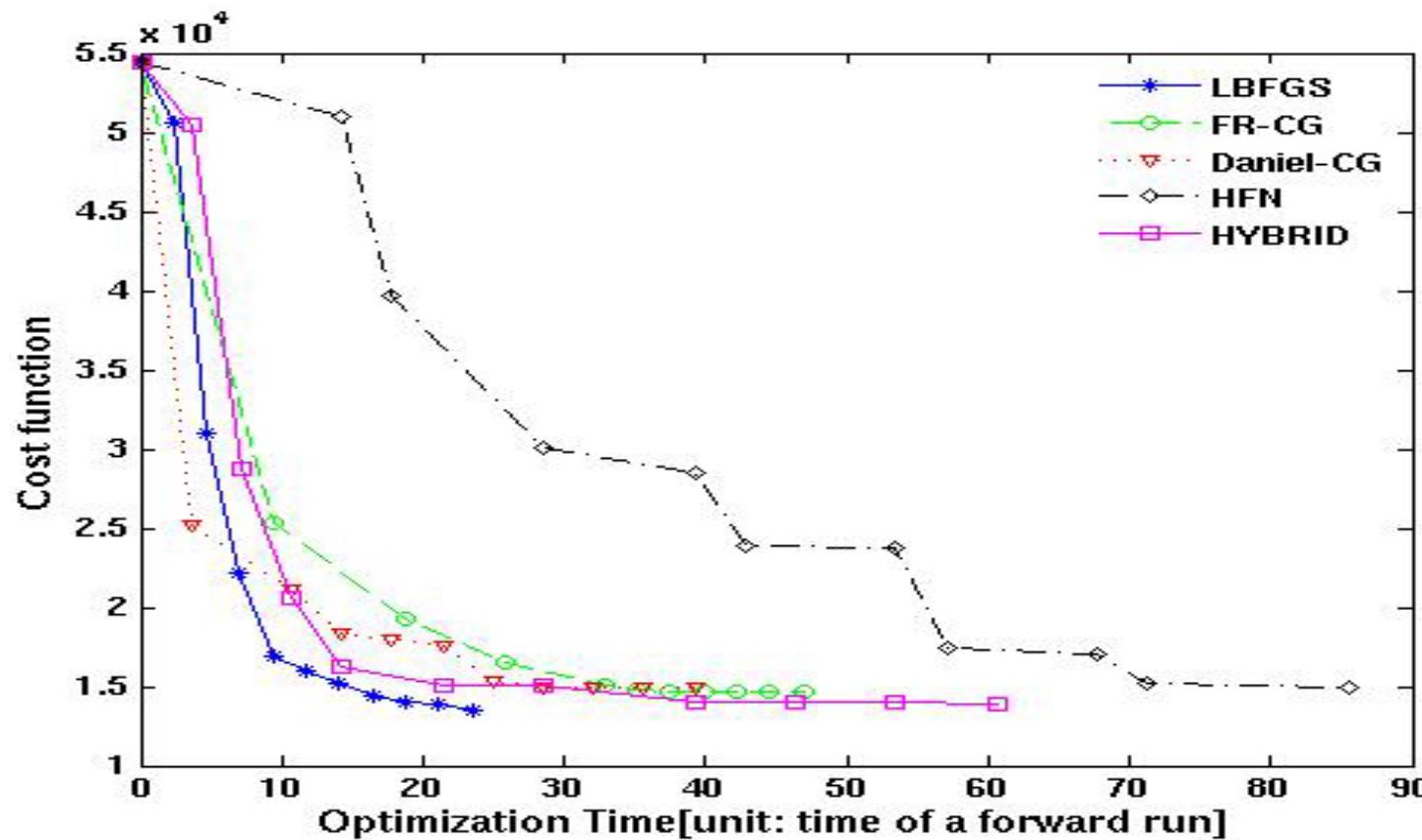
q-q modeled O₃ vs observations

[Chai et al., 2006]



Second order adjoints provide Hessian-vector products useful in optimization and analysis

$$\lambda^0 = \nabla_{y^0} \psi \quad \sigma^0 = (\nabla_{y^0, y^0}^2 \psi) \cdot \delta y^0$$



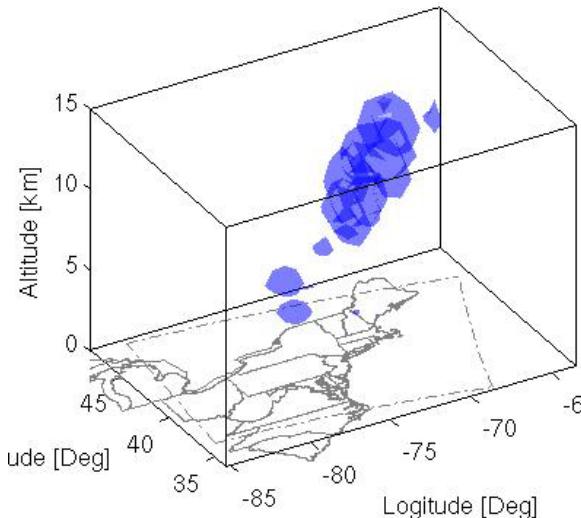
[Zhang and Sandu, 2007]

The smallest Hessian eigenvalues/vectors approx. the principal components of the a posteriori error field

$$\left(\nabla_{\mathbf{y}^0, \mathbf{y}^0}^2 \Psi \right)^{-1} \approx \text{cov}(\mathbf{y}^0)$$

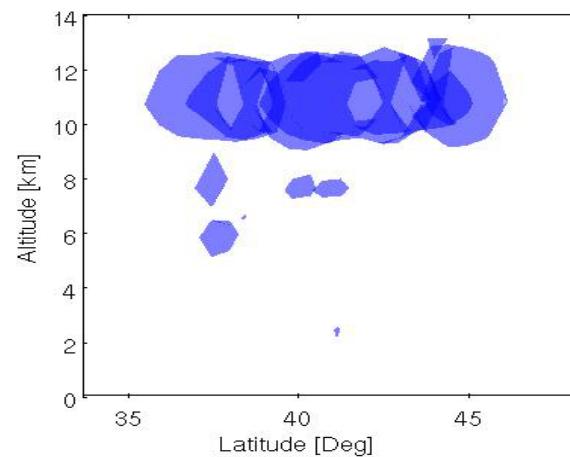
	First	Second	Third	Fourth	Fifth
$\lambda(H)$	7.54e-25	1.15e-23	4.04e-23	8.47e-23	1.42e-22
$\lambda(P)$	1.33e+24	8.70e+22	2.48e+22	1.18e+22	7.04e+21
STD	47 ppb	3 ppb	0.87 ppb	0.41 ppb	0.25 ppb

(a) 3D view (5ppb)

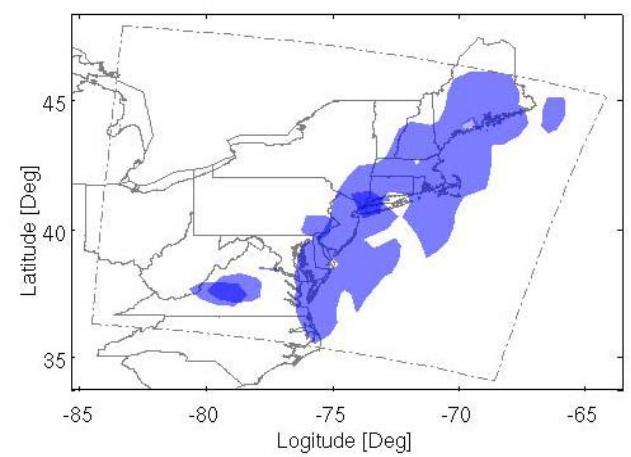


[Zhang and Sandu, 2007]

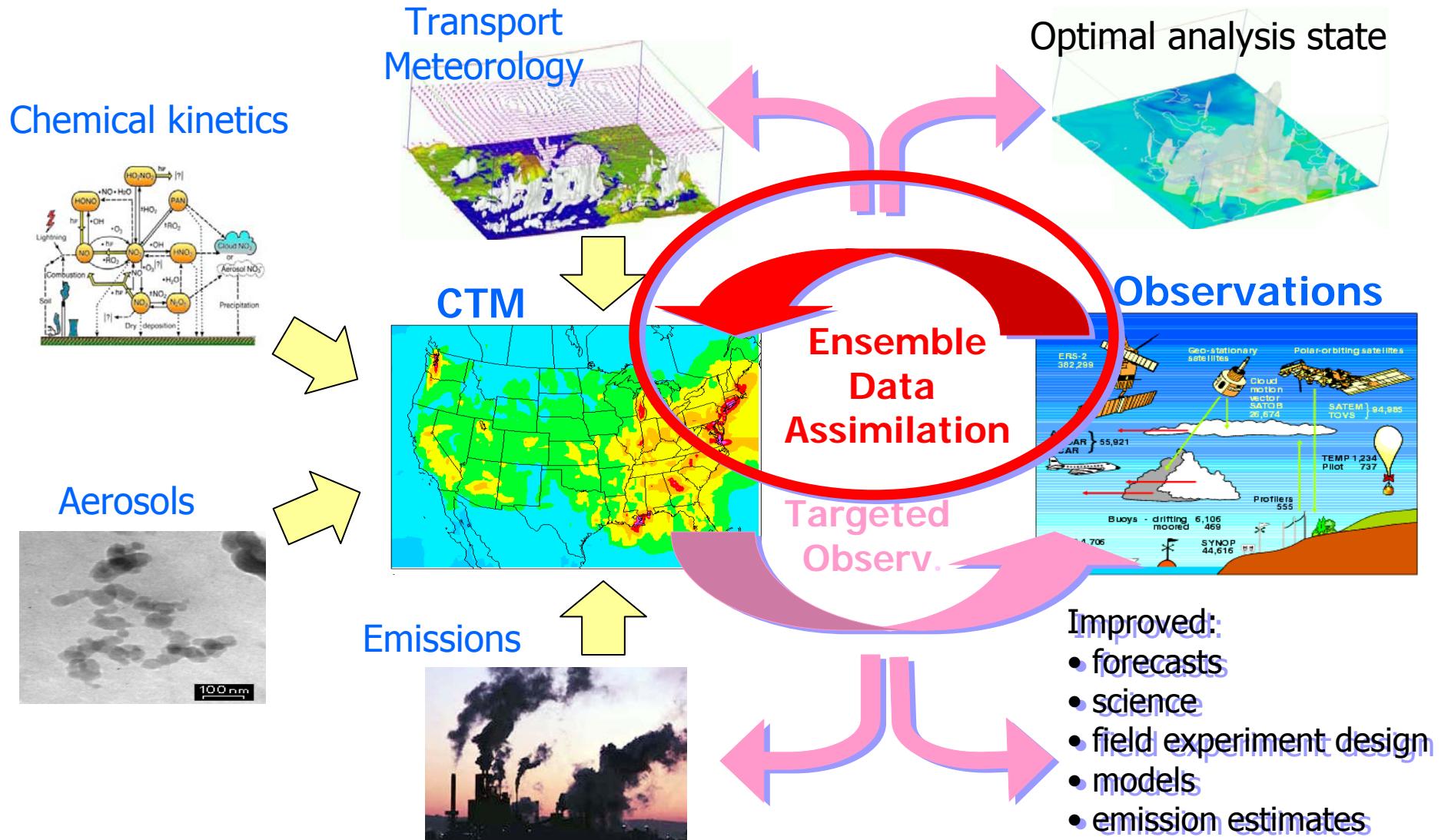
(b) East view



(c) Top view



Ensemble-based chemical data assimilation can complement variational techniques



The Ensemble Kalman Filter (EnKF) popular in NWP but not extensively used before with CTMs

$$\mathbf{y}_f^k = M(t^{k-1}, \mathbf{y}_a^{k-1})$$

$$\mathbf{y}_a^k = \mathbf{y}_f^k + \mathbf{P}_f^k \mathbf{H}_k^T (\mathbf{R}_k + \mathbf{H}_k \mathbf{P}_f^k \mathbf{H}_k^T)^{-1} (\mathbf{z}_{obs}^k - \mathbf{H}_k \mathbf{y}_f^k)$$

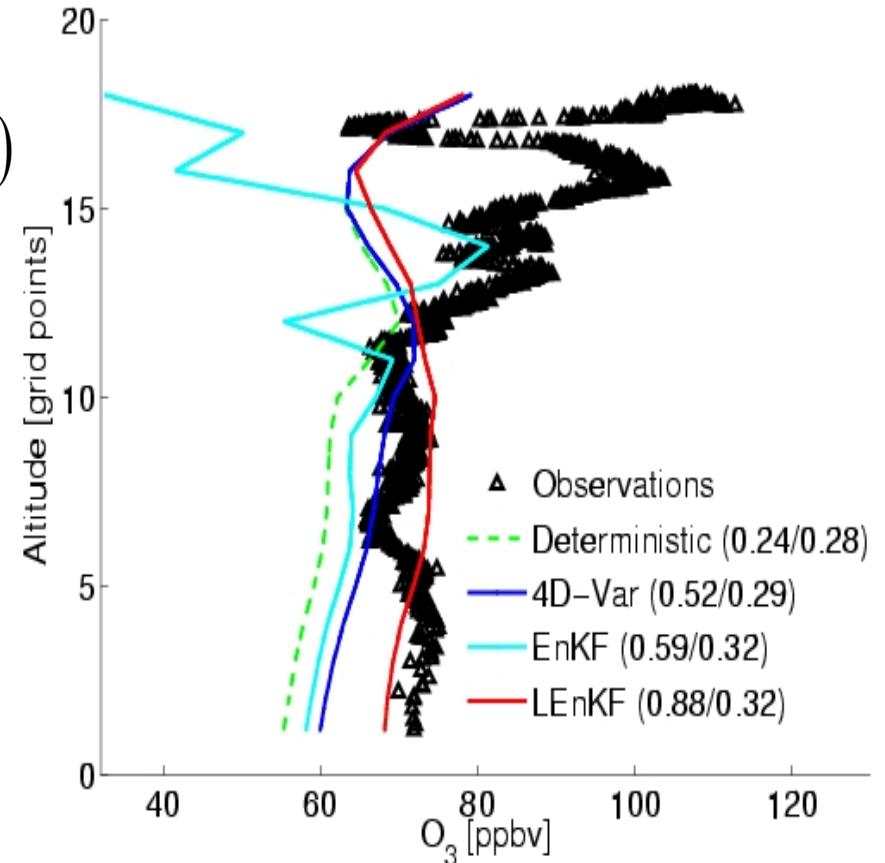
Specify initial ensemble (sample B)

Covariance inflation: Prevents filter divergence (additive, multiplicative, model-specific)

Covariance localization (limit long-distance spurious correlations)

Correction localization (limit increments away from observations)

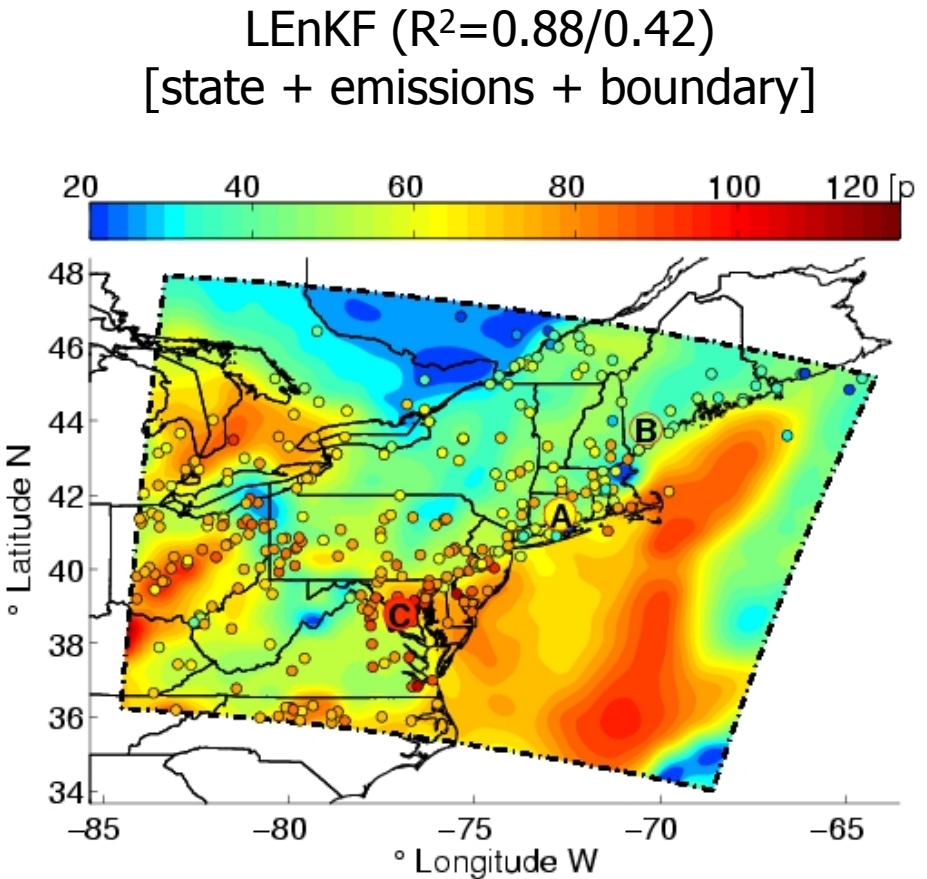
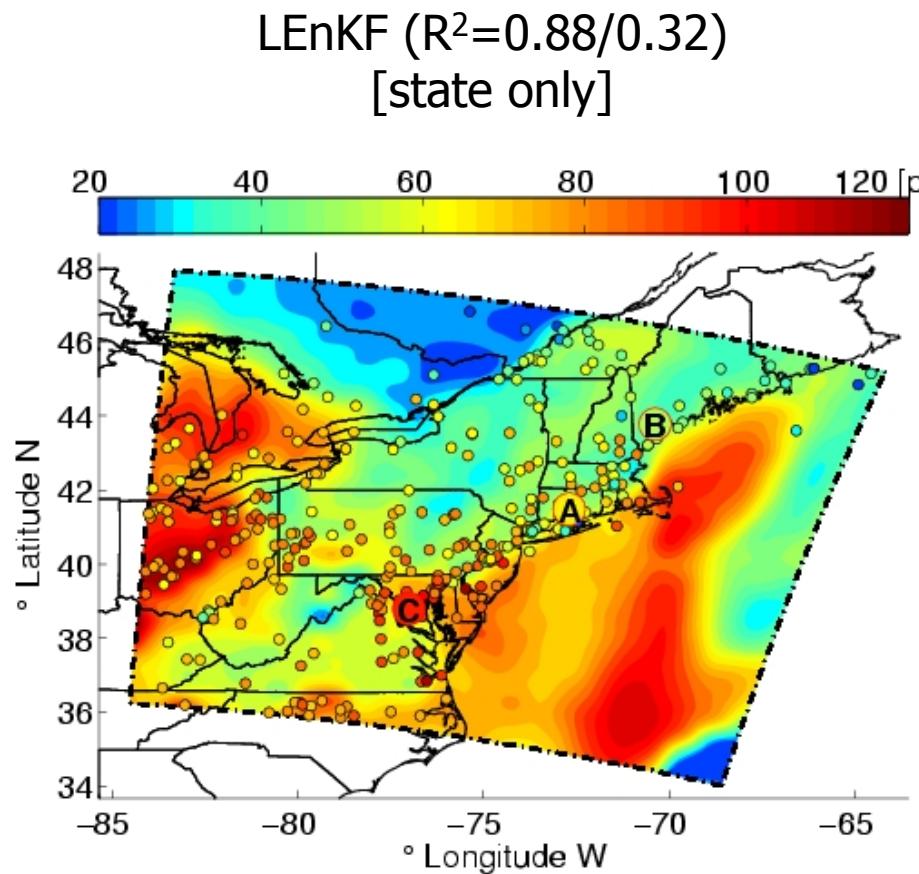
[Constantinescu et al., 2007]



Ozonesonde S2 (18 EDT, July 20, 2004)

LEnKF assimilation of emissions and boundaries together with the state can improve the forecast

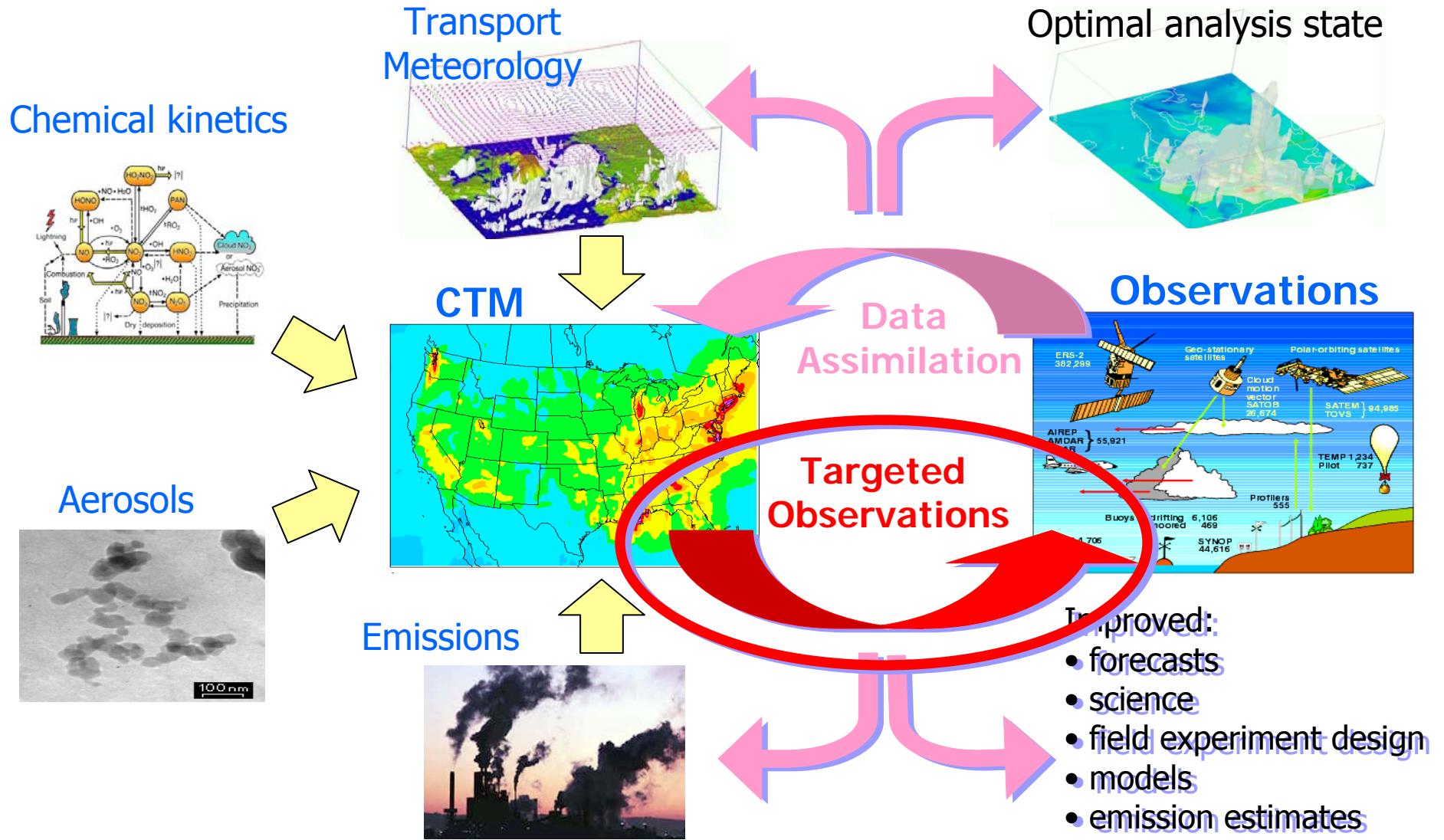
Ground level ozone at 2pm EDT, July 21, 2004 (in forecast window)



Summary

- **Assimilation of chemical observations into CTMs:**
 - enhances reanalysis of fields
 - improves model forecast skills
 - provides top-down estimate of emission inventories
- **Current state of the art:**
 - the tools needed for 4D-Var chemical data assimilation are mature: (adjoints for stiff systems, aerosols, transport; singular vectors, parallelization and multi-level checkpointing schemes, models of background errors)
 - their strengths demonstrated using real (field campaign) data; ambitious science projects are ongoing
 - ensemble-based chemical data assimilation is new, but promising
- **Emerging needs and research directions:**
 - hybrid methods (combining ensemble and variational approaches)
 - second order adjoints and optimization
 - reduced order models

The model can be used to place the observations in the locations of maximum informational benefit



Best observation locations are different for different chemical species

$$T = \sum_{k \geq 1} \frac{\sigma_k^2}{\sigma_{\max}^2} s_k^2$$

(criterion based on SVs)

Verification:
Korea, ground O_3 ,
0 GMT, Mar/4/2001

